Antral follicle count in clinical practice: analyzing clinical relevance

Albert Hsu, M.D., Margaret Arny, Ph.D., Alexander B. Knee, M.S., Carrie Bell, M.D., Elizabeth Cook, Ph.D., Amy L. Novak, R.D.M.S., and Daniel R. Grow, M.D.

Division of Reproductive Endocrinology, Departments of Obstetrics and Gynecology, Baystate Medical Center, Western Campus of Tufts University School of Medicine, Springfield, Massachusetts

Objective: To determine the clinical relevance of obtaining antral follicle counts (AFC) before ovarian stimulation in an IVF program.
Design: Retrospective cohort study.
Setting: An IVF program in a large academic teaching hospital.
Patient(s): A total of 1,049 stimulated IVF cycles in 734 subjects between September 2003 and December 2007 selected from our program's database.
Intervention(s): Basal antral follicles (AFCs) (3 mm–10 mm) were counted via ultrasound scan on cycle day 3 in luteal leuprolide acetate stimulations, or after at least 2 weeks of oral contraceptives in microdose leuprolide acetate stimulations. Patients were grouped according to basal AFC, and outcome parameters compared for AFC groups within each stimulation protocol.
Main Outcome Measure(s): Oocytes retrieved, ovarian response, implantation rate, cancellations, pregnancy, pregnancy loss, and live births per cycle start.
Result(s): Antral follicle count grouping is predictive of threefold change in ovarian response to gonadotropins and oocytes retrieved. Low AFC did predict a higher cancellation rate. Antral follicle count did not predict implantation rate, pregnancy rate, or live birth rate per cycle start.
Conclusion(s): Antral follicle count may be helpful in determining stimulation protocol, as it is the most reliable determinant of oocytes retrieved per starting FSH dose. Antral follicle count predicts ovarian response, not embryo quality or pregnancy. (Fertil Steril® 2011;95:474–9. ©2011 by American Society for Reproductive Medicine.)

Key Words: Antral follicle count, AFC, IVF, ovarian response, pregnancy, IVF cancellation rate, miscarriage, low responder

The ovarian antral follicle count (AFC) has emerged as a useful predictor of ovarian response and stimulation quality in assisted reproductive technologies (ART), and is found more predictive than age or basal serum FSH (1). Many practices have incorporated AFC into the evaluation of patients with infertility for counseling purposes and to determine the gonadotropins dose in ART cycles (2).

Antral follicle count is easy to determine by transvaginal ultrasound examination, has relatively low intercycle variability, and has low to moderate interobserver variability (2–5). Antral follicle count is the best prospective predictor of oocytes retrieved and peak E_2 with ovarian stimulation (5, 6). Repeated measures of AFC in the same patient in different cycles do not markedly improve the predictive value (5). For the same patient, waiting for a cycle with a higher AFC does not improve ovarian stimulation response or oocytes retrieved compared with cycles in the same patient with lower AFC (2).

However, significant questions remain regarding what influences the AFC, and furthermore whether and how AFC affects ART outcomes. Several works have suggested that AFC is predictive of pregnancy outcome, specifically that patients with low AFC had low pregnancy rates, even after controlling for age (7, 8). Others have shown that AFC is predictive of pregnancy in controlled ovarian stimulation–IUI cycles (9). The positive influence of increasing AFC on pregnancy outcome would seem intuitive in light of recent works suggesting that even within age categories, as the number of oocytes retrieved increases, pregnancy rates increase (10). If AFC is the best predictor of oocytes retrieved, then should AFC not also predict pregnancy? However, a recent meta-analysis questions the ability of low AFC to predict nonpregnancy (11), and a recent study in egg donors suggests that AFC does not predict implantation rate for recipients of donor egg ART (12).

Received September 28, 2009; revised March 3, 2010; accepted March 8, 2010; published online April 30, 2010.
A.H. has nothing to disclose. M.A. has nothing to disclose. A.B.K. has nothing to disclose. C.B. has nothing to disclose. E.C. has nothing to disclose. A.L.N. has nothing to disclose. D.R.G. has nothing to disclose.
Supported by the Department of Obstetrics and Gynecology, Baystate Medical Center, the Western Campus of Tufts University School of Medicine, Springfield, Massachusetts.
Presented at the 64th annual meeting of the American Society for Reproductive Medicine, San Francisco, California, November 8–12, 2008.
Reprint requests: Daniel R. Grow, M.D., Department of Obstetrics and Gynecology, Baystate Medical Center, 759 Chestnut St., Springfield, MA 01199 (FAX: 413-784-8166; E-mail: Daniel.Grow@baystatehealth.org).

0015-0282/$36.00
doi:10.1016/j.fertnstert.2010.03.023

Antral follicle count measurements have been a routine part of clinical practice in our IVF program since 2003 and have been collected prospectively in >1,300 IVF cycles. This work represents the largest collection of AFC measurements that we know of in patients with infertility and examines the clinical relevance of AFC measurements in a moderate-sized academic ART program. We hypothesize that prospective determination of AFC will predict oocytes retrieved but not pregnancy outcome.

MATERIALS AND METHODS

After obtaining approval to perform this study from our Institutional Review Board, we conducted the following retrospective cohort study. Basal AFCs were determined by ultrasound scan in a group of 1,049 stimulated IVF cycles in 734 patients between September 2003 and December 2007. Antral follicles (3 mm–10 mm) were counted before initiation of gonadotropins for stimulation. This was performed by a single ultrasonographer and recorded as part of the IVF-ET cycle information that was entered into our program's IVF database.

Stimulation protocols for IVF-ET were determined before a patient started initial stimulation drugs, with protocol determined before the AFC was performed. Starting dose sometimes was adjusted on the basis of the AFC. Patients in our program typically undergo a superovulation-IUI cycle before proceeding to IVF, which helps determine ovarian responsiveness. Normal responders typically use a long GnRH-analogue (GnRH-a) protocol, termed luteal leuprolide acetate (LA), wherein LA (Lupron; TAP Pharmaceuticals Inc., Chicago, IL) was started on day 21 of the luteal phase, 0.5 mg daily until the start of gonadotropins, when the LA dose is decreased to 0.25 mg daily until the day of hCG. The AFC was determined on menstrual day 2 or 3 before the initiation of gonadotropins. During the luteal LA stimulation, the starting dose of gonadotropins was determined by physician discretion but always contained a fixed amount of recombinant FSH (rFSH) supplemented by one ampule (75 IU) of hMG, which continued together for 5 days of stimulation, after which dosing may be altered.

Patients suspected of having a low or poor response to gonadotropins underwent a microdose LA stimulation (13). During the microdose LA stimulation, the AFC was determined typically on cycle day 18 of the oral contraceptive pill pack before the start of ovarian stimulation. This protocol calls for a fixed dose of 450 IU rFSH for the first 5 days of ovarian stimulation with gonadotropins.

It was the protocol of our program to record basal AFC in every patient before starting stimulation. All patients with a recorded AFC using one of the two designated protocols were included in the analysis. There were 729 cycles that used the luteal LA stimulation protocol and 320 cycles that used the microdose LA protocol with AFC data available for the analysis. Parameters of interest included age, body mass index (BMI, calculated as kilograms per square meter), basal (day 2 or 3) FSH, number of oocytes retrieved, ovarian response, implantation rate, cancellation rate, clinical pregnancy, pregnancy loss, and live births. All rates were calculated by cycle start (any cycle in which gonadotropins were initiated). Clinical pregnancy refers to a pregnancy with an intrauterine sac on ultrasound examination. Ovarian response was defined by the number of oocytes retrieved for every 75 IU of FSH administered (the traditional ampule of FSH) as a starting dose of gonadotropins. Ovarian response also was calculated as eggs retrieved/total gonadotropin exposure [(oocytes/total IU rFSH) × 1,000]. Cycle cancellation (no oocyte retrieval and no ET) occurred when there were fewer than three oocytes expected at the time of oocyte retrieval, or for fear of severe ovarian hyperstimulation. Any cause for cancellation was calculated, as well as cancellation for low ovarian response and cancellation for fear of ovarian hyperstimulation. Treatment cycles wherein patients were assigned to other ovarian stimulation protocols were excluded from the analysis.

Patients were stratified by stimulation protocol and then divided into four groups based on AFC, 1–5, 6–10, 11–15, and 16+ antral follicles. Patient and IVF cycle performance characteristics then were analyzed with respect to AFC group.

Data were extracted from the IVF database, deidentified, and analyzed with use of Stata 10.1 (StataCorp LP, College Station, TX). Statistical significance for all tests was set at an alpha ≤ 0.05. We presented numbers and percentages for categorical data and medians and interquartile ranges for continuous variables, which were not normally distributed. The only exception to this was implantation rate, where we presented means and SDs because we believed that this was more clinically meaningful. To determine significant differences for age, BMI, and day 3 FSH across the AFC distributions we used the Kruskal-Wallis equality-of-populations rank test. Significant differences for infertility diagnosis by AFC category were calculated with use of a χ^2 test. Nonparametric tests for trend across AFC categories were conducted for all continuous outcomes. For all binary outcomes, we calculated an age-adjusted odds ratio (OR) and 95% confidence interval (CI) using multivariate logistic regression. These models used a clustered sandwich estimator to adjust for subjects with multiple cycles. For these binary outcomes, a P value for trend was calculated with use of logistic regression to evaluate the significance of a monotonic association between AFC category and the respective outcome.

RESULTS

There were 729 (69.5%) cycles reviewed in patients undergoing the luteal LA stimulation protocol and 320 (30.5%) cycles in patients undergoing the microdose LA protocol. Patients undergoing an ET are listed, with a statistically higher percentage of patients reaching ET as the AFC increases. The lowest AFC group had 85.7% and 66.7% of patients reaching ET in the luteal LA and microdose LA groups respectively, compared with 92.7% and 84% for the groups with ≥16 AFC. Within stimulation groups, there were minimal differences in age and BMI when different AFC groups were compared, though these differences were statistically significant (Table 1). Low-AFC patients were older in the luteal LA stimulation patients, displaying the expected age-related decline in AFC within the luteal LA group. This relationship of declining AFC with advancing age did not hold in the microdose LA group, suggesting that there are other determinants of AFC than age. There was a difference in mean basal FSH between AFC groups in the luteal LA patients when comparing those with <10 antral follicles with those with >10 antral follicles ($P<.0001$). As expected, those with a higher AFC had a slightly lower FSH. Basal FSH levels were not different between AFC groups in the microdose LA stimulation patients.

Oocytes retrieved for each AFC group are listed in Table 2. The interquartile range for the oocytes retrieved overlaps with each of the three higher-AFC groups, and the trend toward higher egg count with each AFC group is apparent and statistically relevant. (If cancelled cycles were included there likely would be a lower number of oocytes retrieved in the lowest AFC group.) The response of the ovary to gonadotropins (defined here as the number of eggs retrieved per 75 IU of FSH starting dose maintained for 5 days) is predicted with the AFC group. Among the luteal LA stimulations,

TABLE 1

Patient and cycle characteristics by stimulation protocol and AFC category.

	Subjects (n)	Subject cycles, n (%)[a]	ET, n (%)[a]	Age (y)[b]	BMI[b,c,e]	Day 3 FSH (IU/mL)[b,d,f]
Luteal LA protocol						
Overall	550	729 (100)	659 (90.4)	34 (30–38)	26 (23–31)	6 (5–7)
AFC 1–5	20	21 (2.9)	18 (85.7)	37 (32–40)	26 (21–33)	7 (5–8)
AFC 6–10	140	159 (21.8)	129 (81.1)	35 (31–38)	24 (22–29)	7 (6–8)
AFC 11–15	187	209 (28.7)	197 (94.3)	34 (32–37)	25 (22–30)	6 (5–8)
AFC 16+	272	340 (46.6)	315 (92.7)	33 (30–37)	26 (23–32)	6 (5–7)
P value	—	—	.001	.0506	.0321	.0001
Microdose LA protocol						
Overall	245	320 (100)	258 (80.6)	39 (36–41)	25 (22–30)	8 (6–10)
AFC 1–5	68	72 (22.5)	48 (66.7)	39 (36–41)	25 (23–29)	8 (6–11)
AFC 6–10	127	144 (45.0)	119 (82.6)	38 (36–41)	25 (22–30)	8 (6–10)
AFC 11–15	71	79 (24.7)	70 (88.6)	40 (37–42)	25 (22–30)	8 (6–10)
AFC 16+	23	25 (7.8)	21 (84.0)	41 (39–42)	27 (23–30)	8 (6–10)
P value	—	—	.011	.0107	.9823	.8050

[a] *P* value (if present) represents *P* trend calculated from logistic regression model.
[b] Expressed as median and interquartile range (rounded); *P* value calculated with use of Kruskal-Wallis equality-of-populations rank test.
[c] Not all data available for luteal LA BMI: AFC 1–5 (n = 11), AFC 6–10 (n = 133), AFC 11–15 (n = 163), AFC 16+ (n = 262).
[d] Not all data available for luteal LA day 3 FSH: AFC 1–5 (n = 15), AFC 6–10 (n = 119), AFC 11–15 (n = 155), AFC 16+ (n = 256).
[e] Not all data available for microdose LA BMI: AFC 1–5 (n = 52), AFC 6–10 (n = 118), AFC 11–15 (n = 61), AFC 16+ (n = 19).
[f] Not all data available for microdose LA day 3 FSH: AFC 1–5 (n = 63), AFC 6–10 (n = 128), AFC 11–15 (n = 70), AFC 16+ (n = 24).

Hsu. Controversy. Fertil Steril 2011.

there is a greater than threefold difference in ovarian response between the lowest and highest AFC groups, from 1.6 to 5.8 eggs per ampule of gonadotropin. Interquartile ranges for ovarian responsiveness do not overlap between the lowest and highest AFC groups (Fig. 1). Ovarian response for the microdose LA stimulations shows a similar trend with a greater than twofold difference between the lowest and highest AFC groups. When looking at those patients cancelled for low ovarian response, the odds ratio (95% CI) for cancellation with AFC 1–5 and for AFC 6–10 was 8.3 (1.9–35.6) and 7.1 (3.0–16.7), respectively, as compared with the referent group AFC 16+ when using the luteal LA stimulation protocol. The OR for cancellation also was increased in the AFC 1–5 group with use of microdose LA stimulation, 7.4 (2.1–26.9), and significantly higher than the referent group. These low-AFC groups had cancellation rates of 13.3% to 18.3%, compared with the referent groups of 0% to 2.1% (Table 2).

The implantation rates for different AFC groups were similar when using the same stimulation protocol (Table 3). The mean ± SD implantation rate for each AFC group using the luteal LA stimulation protocol was 0.4 ± 0.4. The implantation rate for the AFC groups using microdose LA was lower than for the younger luteal LA patients with a mean of 0.2 ± 0.3, and there was no significant difference between AFC groups. A low AFC did not predict a poorer quality embryo.

Likewise, the pregnancy rate, pregnancy loss rate, and live birth rates per cycle start did not differ between AFC groups for either the luteal LA or the microdose LA stimulation protocols. Even though the cancellation rate was higher in the low AFC groups, the live birth rates per cycle start were not statistically lower, with the live birth rates per cycle start ranging from 42.9% in the AFC 1–5 group to 47.4% in the AFC 11–15 group for luteal LA and AFC 1–5 = 22.2% to AFC 11–15 = 26.6% for the microdose LA cycles.

To evaluate whether our results were influenced by infertility diagnosis, for each stimulation protocol we determined the associations with AFC category and each of our categorical outcomes. For our data regarding luteal LA, we did find some significant associations with overall infertility diagnosis and AFC category ($P<.001$). However, we did not find any differences between infertility diagnosis and cancellation ($P=.112$), clinical pregnancy ($P=.928$), live birth rate ($P=.890$), or pregnancy loss ($P=.902$).

DISCUSSION

This study encompasses 1,049 IVF stimulation cycles in patients undergoing two commonly used stimulation protocols, namely the long GnRH-a down-regulation protocol (luteal LA) and a protocol used for low-responder patients using microdose LA, and represents one of the largest analyses of AFC with respect to ART outcomes that we have encountered.

We have confirmed once again that AFC is more predictive of ovarian response and oocytes retrieved than other available information including age and basal FSH. There was a threefold difference in the number of oocytes retrieved per ampule (75 IU) FSH administered as a starting dose between the lowest and highest AFC group. The increase in ovarian responsiveness was incremental with increasing AFC group within a given stimulation protocol. The response rate for the luteal LA patients was higher than for the microdose LA protocol patients. This difference likely is accounted for partially by the higher mean age in the microdose LA group and by the fact that their physician directed them to the microdose LA protocol because they were known to be poor responders from prior FSH-IUI experience. There were no predetermined strict selection criteria set for assignment of stimulation protocol. Whether a patient's treatment cycle included luteal LA or microdose LA was left solely at the discretion of the treating physician. The stated goal of ovarian stimulation was to achieve between 10 and 20 mature oocytes at retrieval, with stimulation protocol chosen accordingly.

This calculated "response rate" for ovarian responsiveness is perhaps novel and potentially a useful way to characterize the poor responder patient. The "poor responder patient" has been

TABLE 2

Oocytes retrieved, ovarian response, and cancellation rates by stimulation protocol and AFC category.

Group	Oocytes retrieved[a]	Ovarian response: oocytes per ampule of FSH[a]	Ovarian response: (oocytes per total FSH) × 1000[a]	Cancellation: no retrieval and no ET		Cancellation: fewer than three oocytes expected		Cancellation: elevated risk of OHSS	
				n (%)	OR (95% CI)[b]	n (%)	OR (95% CI)[b]	n (%)	OR (95% CI)[b]
Luteal LA protocol (n = 729 cycles)									
No. (%)	682 (93.6)	680 (93.3)	670 (91.9)	70 (9.6)		36 (5.0)[c]		4 (0.6)[c]	
AFC 1–5	10 (7–16)	1.6 (1.2–3.7)	2.9 (1.5–4.4)	3 (14.3)	2.0 (0.5–7.7)	3 (15.8)	8.3 (1.9–35.6)	—	—
AFC 6–10	12 (9–19)	2.7 (1.7–3.7)	3.9 (2.7–6.1)	30 (18.9)	2.9 (1.6–5.2)	21 (13.3)	7.1 (3.0–16.7)	—	—
AFC 11–15	16 (11–22)	3.6 (2.3–5.0)	5.7 (3.8–8.4)	12 (5.7)	0.8 (0.4–1.5)	5 (2.4)	1.2 (0.4–3.7)	—	—
AFC 16+	20 (15–26)	5.8 (4.2–8.3)	9.6 (6.7–15.0)	25 (7.4)	1.0 (Referent)	7 (2.1)	1.0 (Referent)	4 (1.2)	—
P trend	<.001[d]	<.001[d]	<.001[d]	—	<.001[e]	—	<.001[e]	—	—
Microdose LA protocol (n = 320 cycles)									
No. (%)	279 (87.2)	279 (86.4)	279 (86.4)	62 (19.4)		28 (8.9)[f]		3 (1.0)[f]	
AFC 1–5	7 (5–12)	1.2 (0.8–1.8)	1.8 (1.3–3.3)	24 (33.3)	2.7 (0.9–8.2)	13 (18.3)	7.4 (2.1–26.9)	1 (1.4)	—
AFC 6–10	9 (6–13)	1.5 (1.0–2.2)	2.4 (1.6–3.5)	25 (17.4)	1.1 (0.4–3.4)	12 (8.6)	3.1 (0.9–11.3)	—	—
AFC 11–15	12 (8–18)	2.0 (1.3–3.0)	3.4 (2.1–4.7)	9 (11.4)	0.7 (0.2–2.1)	3 (3.9)	1.0 (Referent)	—	—
AFC 16+	16 (11–22)	2.8 (2.0–3.8)	4.9 (3.2–5.8)	4 (16.0)	1.0 (Referent)	0 (0.0)	—	2 (8.0)	—
P trend	<.001[d]	<.001[d]	<.001[d]	—	.011[e]	—	.001[e]	—	—

Note: OHHS = ovarian hyperstimulation syndrome.

[a] Expressed as median and interquartile range.

[b] Age-adjusted OR and 95% CI calculated with use of multivariate logistic regression with clustered sandwich estimator.

[c] Not all data available for luteal LA peak E_2: AFC 1–5 (n = 19), AFC 6–10 (n = 158), AFC 11–15 (n = 206), AFC 16+ (n = 337).

[d] *P* value calculated from nonparametric test for trend across ordered groups.

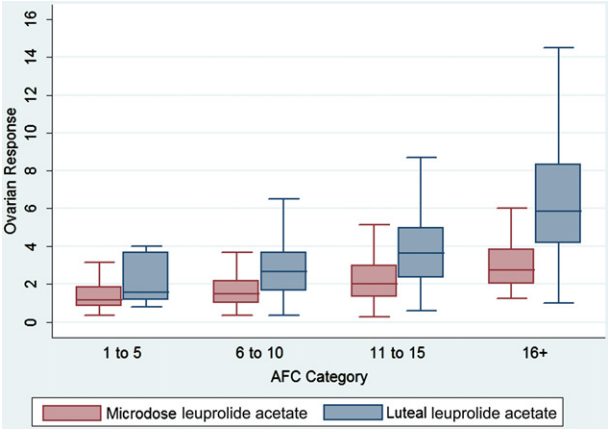
[e] *P* trend calculated from logistic regression model evaluating significance of monotonic association between AFC category and outcome.

[f] Not all data available for microdose LA peak E_2: AFC 1–5 (n = 71), AFC 6–10 (n = 140), AFC 11–15 (n = 78), AFC 16+ (n = 25).

Hsu. Controversy. Fertil Steril 2011.

FIGURE 1

Box and whisker plot of ovarian response by simulation protocol and AFC category. Extreme values outside the 10th and 90th percentiles are not shown.



Hsu. Controversy. Fertil Steril 2011.

referred to in our literature countless times without a precise definition (14). Patients who produce fewer than two oocytes per 75 IU of FSH and who have <10 antral follicles may reasonably define this group. This will be the focus of further study.

One of the earliest references to the AFC was by Ruess et al. in 1996 (15), who suggested that ultrasonographically derived counts of follicles provide a measure of reproductive age that may help to predict age-related phenomena. This work suggests that AFC is a marker of ovarian response with predictive ability that is different than age. Age is associated with both ovarian response and a poorer implantation rate of embryos (11). Antral follice count has no predictive value for implantation rate (embryo quality) in this study but is a much more robust predictor of ovarian response and cancellation rate than is age.

The AFC is easy to determine, and, as suggested by the earliest clinical work with ART, AFC is predictive of ovarian response (7). We agree with the sentinel work of Chang et al. (7), that AFC is easy to determine, <1 minute per ovary, and that this is an important indicator of ovarian response. Several prior publications have suggested a 30% difference in pregnancy rates between the low and higher AFC groups (1, 7, 16). Low AFC groups typically refer to patients with four or fewer AFC. However, as >90% of the microdose LA stimulation patients had AFC counts of <15, we constructed four AFC groups for the analysis, with three groups for those with AFC ≤15 and one for those with AFC >15.

Prospectively collecting AFC in more than 1,000 IVF-ET cycles has allowed us to determine that patients with low AFC are able to achieve pregnancy at a high rate and that AFC does not predict the pregnancy potential of a cycle, after controlling for the stimulation protocol that a patient receives. Implantation rates for the four AFC groups were not different within a given stimulation protocol. The ORs for clinical pregnancy and live birth were also similar between groups. Confidence intervals were relatively wide, particularly in the microdose LA group because of the smaller number of pregnancies and smaller sample size.

Antral follicle count does predict cancellation rate, with a fivefold increase in cancellation rate between the lowest and highest AFC groups as suggested by others (17). Knowing AFC does allow for optimization of stimulation protocol. Knowing the AFC undoubtedly influenced the physicians in our program and may account for the relatively low cancellation rate of <20% in this study, even in the low-AFC groups, compared with other reports with cancellation rates >40% with low-AFC groups (7, 12, 16).

TABLE 3

Pregnancy outcome rates and associated risk by stimulation protocol and AFC category.

		Clinical pregnancy		Pregnancy loss		Live births	
Group	**Implantation rate**	**n (%)**	**OR (95% CI)[a]**	**n (%)**	**OR (95% CI)[a]**	**n (%)**	**OR (95% CI)[a]**
Luteal LA protocol (n = 729 cycles)							
No. (%)	659 (90.4)	375 (51.4)		52 (7.1)		323 (44.3)	
AFC 1–5, mean (SD)	0.4 (0.4)	10 (47.6)	0.8 (0.4–2.0)	1 (4.8)	0.6 (0.1–5.1)	9 (42.9)	0.9 (0.4–2.3)
AFC 6–10, mean (SD)	0.4 (0.4)	68 (42.8)	0.7 (0.5–1.0)	11 (6.9)	1.0 (0.5–2.2)	57 (35.9)	0.7 (0.4–1.0)
AFC 11–15, mean (SD)	0.4 (0.4)	117 (56.0)	1.1 (0.8–1.6)	18 (8.6)	1.4 (0.7–2.6)	99 (47.4)	1.1 (0.8–1.5)
AFC 16+, mean (SD)	0.4 (0.4)	180 (53.0)	1.0 (Referent)	22 (6.5)	1.0 (Referent)	158 (46.5)	1.0 (Referent)
P trend	0.116[b]	—	0.092[c]	—	0.869[c]	—	0.077[c]
Microdose LA protocol (n = 320 cycles)							
No. (%)	258 (80.6)	106 (33.1)		29 (9.1)		77 (24.1)	
AFC 1–5, mean (SD)	0.2 (0.3)	22 (30.6)	1.1 (0.4–3.0)	6 (8.3)	0.7 (0.2–2.9)	16 (22.2)	1.4 (0.4–4.7)
AFC 6–10, mean (SD)	0.2 (0.4)	46 (31.9)	1.1 (0.4–3.1)	10 (6.9)	0.6 (0.2–2.2)	36 (25.0)	1.6 (0.5–5.1)
AFC 11–15, mean (SD)	0.2 (0.3)	31 (39.2)	1.6 (0.6–4.6)	10 (12.7)	1.1 (0.3–4.0)	21 (26.6)	1.8 (0.5–6.2)
AFC 16+, mean (SD)	0.2 (0.3)	7 (28.0)	1.0 (Referent)	3 (12.0)	1.0 (Referent)	4 (16.0)	1.0 (Referent)
P trend	0.116[b]	—	0.578[c]	—	0.313[c]	—	0.913[c]

[a] Age-adjusted OR and 95% CI calculated with use of multivariate logistic regression with clustered sandwich estimator.
[b] *P* value calculated from nonparametric test for trend across ordered groups.
[c] *P* trend calculated from logistic regression model evaluating significance of monotonic association between AFC category and outcome.

Hsu. Controversy. Fertil Steril 2011.

Contrary to early studies of AFC, which described a correlation between AFC and pregnancy outcome, there is a growing body of work suggesting otherwise. A recent meta-analysis that produced receiver-operator curves for 13 AFC studies did not find a relationship between low AFC and pregnancy (11, 18). A recent very large study examining 975 oocyte recipient cycles suggested that AFC cannot be used to predict oocyte or embryo quality or IVF outcome when donor eggs are used (12). Together these works suggest that, unlike age, AFC is a pure marker of ovarian response and does not predict oocyte or embryo developmental competence or potential for pregnancy. Whereas there are ages that predict very low implantation rates, we cannot say the same for AFC.

Within the luteal LA stimulation group there were significant differences among some patient characteristics within the AFC groups. Age ranged from a median of 33 to 37 years between the highest and lowest AFC, as would be clinically predicted. Day 3 basal FSH levels ranged from a median of 6 to 7 IU/mL between the highest and lowest AFC groups. Though this difference was statistically significant, it represents a narrow range that is hard to equate to clinical significance. Body mass index differences between AFC groups were significant but did not display clinically discernable trends.

There was a significant age difference between AFC groups for the microdose LA patients with medians ranging from 38 to 41 years. Surprisingly, the higher-AFC groups had older patients, which may suggest some provider bias, placing older patients into the more aggressive microdose LA stimulation protocol without regard for the AFC information available.

This study was not limited to patients doing only a single IVF cycle, and, as multiple cycles were performed in some patients, selection bias may have been introduced. Furthermore, we know that pregnancy outcome is influenced somewhat by the infertility diagnosis. We found that the distributions of infertility diagnoses between AFC groups were not similar. This is not surprising as low-AFC groups may have a higher likelihood of diminished ovarian reserve, whereas high-AFC groups may contain patients with other forms of ovarian dysfunction such as polycystic ovary syndrome. For both drug protocols, we determined whether the ORs were similar over different infertility diagnoses and found no reasonable differences across strata; however, this may be due to the small numbers within certain diagnoses in our dataset. To account for differences in the frequency of diagnoses within AFC groups, we included infertility diagnosis in our adjusted model to determine whether there were significant changes in the adjusted OR and found minimal changes in the ORs or the CIs. Therefore, we feel confident that the age-adjusted ORs that we have reported are reasonable estimates of risk related to our outcomes.

Though our program occasionally uses stimulation protocols besides those analyzed herein, luteal LA and microdose LA, the numbers of patients in other stimulation protocol groups were relatively small, and excluded from this analysis. There were insufficient numbers to make any conclusions when the stimulation groups were analyzed separately.

In summary, ultrasound determination of AFC is a useful technique for the prediction of ovarian response and may help predict those patients at risk of producing three or fewer oocytes for retrieval and cycle cancellation. Among those patients who do go on to oocyte retrieval, AFC did not predict the chance of pregnancy, pregnancy loss, or live birth. Antral follicle count was the most sensitive predictor of ovarian response to gonadotropin.

REFERENCES

1. Nahum R, Shifren JL, Chang Y, Leykin L, Isaacson K, Toth TL. Antral follicle assessment as a tool for predicting outcome in IVF—is it a better predictor than age and FSH? J Assist Reprod Genet 2001;18:151–5.
2. Hansen KR, Morris JL, Thyer AC, Soules MR. Reproductive aging and variability in the ovarian antral follicle count: application in the clinical setting. Fertil Steril 2003;80:577–83.
3. Pache TD, Wladimiroff JW, DeJong FH, Hop WC, Fauser BC. Growth patterns of nondominant ovarian follicles during the normal menstrual cycle. Fertil Steril 1990;54:638–42.
4. Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW, TeVelde ER. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. Fertil Steril 1999;72:845–51.
5. Bancsi LF, Broekmans FJ, Eijkemans MJ, DeJong FH, Habbema JD, TeVelde ER. Predictors of poor ovarian response in in-vitro fertilization: a prospective study comparing basal markers of ovarian reserve. Fertil Steril 2002;77:328–36.
6. Ng EH, Tang OS, Ho PC. The significance of the number of antral follicles prior to stimulation in predicting ovarian responses in an IVF programme. Hum Reprod 2000;15:1937–42.
7. Chang MY, Chiang CH, Hsieth TT, Soong YK, Hsu KH. Use of the antral follicle count to predict the outcome of assisted reproductive technologies. Fertil Steril 1998;69:505–10.
8. Muttukrishna S, McGarrigle H, Wakim R, Khadum I, Ranieri DM, Serhal P. Antral follicle count, antimullerian hormone and inhibin B: predictors of ovarian response in assisted reproductive technology? Br J Obstet Gynaecol 2005;112:1384–90.
9. Erdem A, Erdem M, Biberoglu K, Hayit O, Arslan M, Gursoy R. Age-related changes in ovarian volume, antral follicle counts and basal FSH in women with normal reproductive health. J Reprod Med 2002;47:835–9.
10. Yih MC, Spandorfer SD, Rosenwaks Z. Egg production predicts a doubling of in vitro fertilization pregnancy rates even within defined age and ovarian reserve categories. Fertil Steril 2005;83:24–9.
11. Broekmans F, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update 2006;12:685–718.
12. Melo MA, Garrido N, Alvarez C, Bellver J, Mesequer M, Pellicer A, et al. Antral follicle count (AFC) can be used in the prediction of ovarian response but cannot predict the oocyte/embryo quality or the in vitro fertilization outcome in an egg donation program. Fertil Steril 2009;91:148–56.
13. Schoolcraft W, Schlenker T, Gee M, Stevens J, Wagley L. Improved controlled ovarian hyperstimulation in poor responder in vitro fertilization patients with a microdose follicle-stimulating hormone flare, growth hormone protocol. Fertil Steril 1997;67:93–7.
14. Loutradis D, Vomvolaki E, Drakakis P. Poor responder protocols for in-vitro fertilization: options and results. Curr Opin Obstet Gynecol 2008;20:374–8.
15. Ruess ML, Kline J, Santos R, Levin B, Timor-Tritsch I. Age and the ovarian follicle pool assessed with transvaginal ultrasonography. Am J Obstet Gynecol 1996;174:624–7.
16. Frattarelli JL, Levi AJ, Miller BT, Segars JH. A prospective assessment of the predictive value of basal antral follicles in in vitro fertilization cycles. Fertil Steril 2003;80:350–5.
17. Frattarelli JL, Lauria-Costab DF, Miller BT, Bergh PA, Scott RT. Basal antral follicle number and mean ovarian diameter predict cycle cancellation and ovarian responsiveness in assisted reproductive technology cycles. Fertil Steril 2000;74:512–7.
18. Hendriks DJ, Mol BW, Bancsi LF, TeVelde ER, Broekmans FJ. Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: a meta-analysis and comparison with basal follicle-stimulating hormone level. Fertil Steril 2005;83:291–301.